RUN PROBABILITIES IN SEQUENCES OF BERNOULLI TRIALS

Steven J. Schwager

BU-755-M*

September 1981

Abstract

The probability that a sequence of n Bernoulli trials contains a run of at least k successive 1's is found as a function of n, k, and the probability of a 1 at any trial, which may vary with the trial. Applications are discussed, and computer calculations examine the accuracy of an approximation given by Feller. The method of using Bonferroni inequalities to obtain upper and lower bounds to the probability of a run of k 1's is shown to behave poorly; upper and lower bounds are not close to the true probability and successive bounds diverge in many cases.

RUN PROBABILITIES IN SEQUENCES OF BERNOULLI TRIALS

Steven J. Schwager

BU-755-M*

September 1981

1. Statement of the problem. Let $X_1, \cdots, X_n$ be a sequence of Bernoulli trials, that is, a sequence of independent, identically distributed dichotomous random variables, each taking the value 1 with probability p and 0 with probability $1-p$. It is desired to find the probability that the sequence of n Bernoulli trials contains a run of at least k successive 1's.

This question arose in connection with interpretation of radar astronomical observations of a minor planet (Ostro et al., 1980). The data were echo power spectral density estimates at each of 433 Doppler frequencies separated by 1.22 Hertz. The empirically determined background filter shape had been removed from the raw spectrum, and the resultant background-free spectrum had been normalized to the root-mean-square fluctuation in the receiver noise. If no echo were present, it was known from both a priori theoretical considerations and a posteriori experimental evidence that the spectral estimates would be expected to behave as realizations of a sequence of independent, identically distributed Normal(0,1) random variables. However, if the target had a sufficiently large radar cross section, a radar echo would be expected to produce a sequence of above average readings in some portion of the frequency band. The 530-Hertz ($=433 \times 1.22$) frequency band was centered on the a priori Doppler frequency of the echo. This suggested that a test for the presence of an echo could be

---

* In the Biometrics Unit Mimeo Series, Cornell University, Ithaca, New York 14853.

based on the length of the longest run of positive readings. The probability of a run of k positive readings out of n under no-echo or null conditions, when the observations are a random sample from a Normal(0,1) distribution, gives the observed significance level of this test.

As shown in Figure 1, the radar echo spectrum contains a suspiciously long run of $k = 28$ positive readings occurring in a sequence of $n = 433$ frequencies. None of these 28 successive positive amplitudes has a value less than .3 . This example from radar astronomy and others like it suggested that it was useful to know how long a run constituted significant evidence of a departure from null conditions.

Knowledge of the probability of a run of at least k successive 1's can be useful in many practical situations. Tests based on runs, and in particular on the length of the longest run, were mentioned by Gibbons (1971):

> "Since a run which is unusually long reflects a tendency for like objects to cluster and therefore possibly a trend, Mosteller (1941) has suggested a test for randomness based on the length of the longest run. Exact and asymptotic probability distributions of the numbers of runs of given length are discussed in Mood (1940)."

In much work on runs, such as Shaughnessy (1981), the number of 1's and 0's occurring in the data is assumed known and is used for conditioning. This conditional approach is not needed when the underlying probability of obtaining a 1 is known.

Feller (1968, Section XIII.7) obtained generating functions related to the probability of a run of at least k 1's in a sequence of n Bernoulli trials. He used an unusual definition of a run to make runs recurrent events. The probability of a run of at least k 1's can be computed in principle from one of these generating functions, but the labor involved is very great. Perhaps because of this, Feller provides the approximation for the probability of no runs of length k:

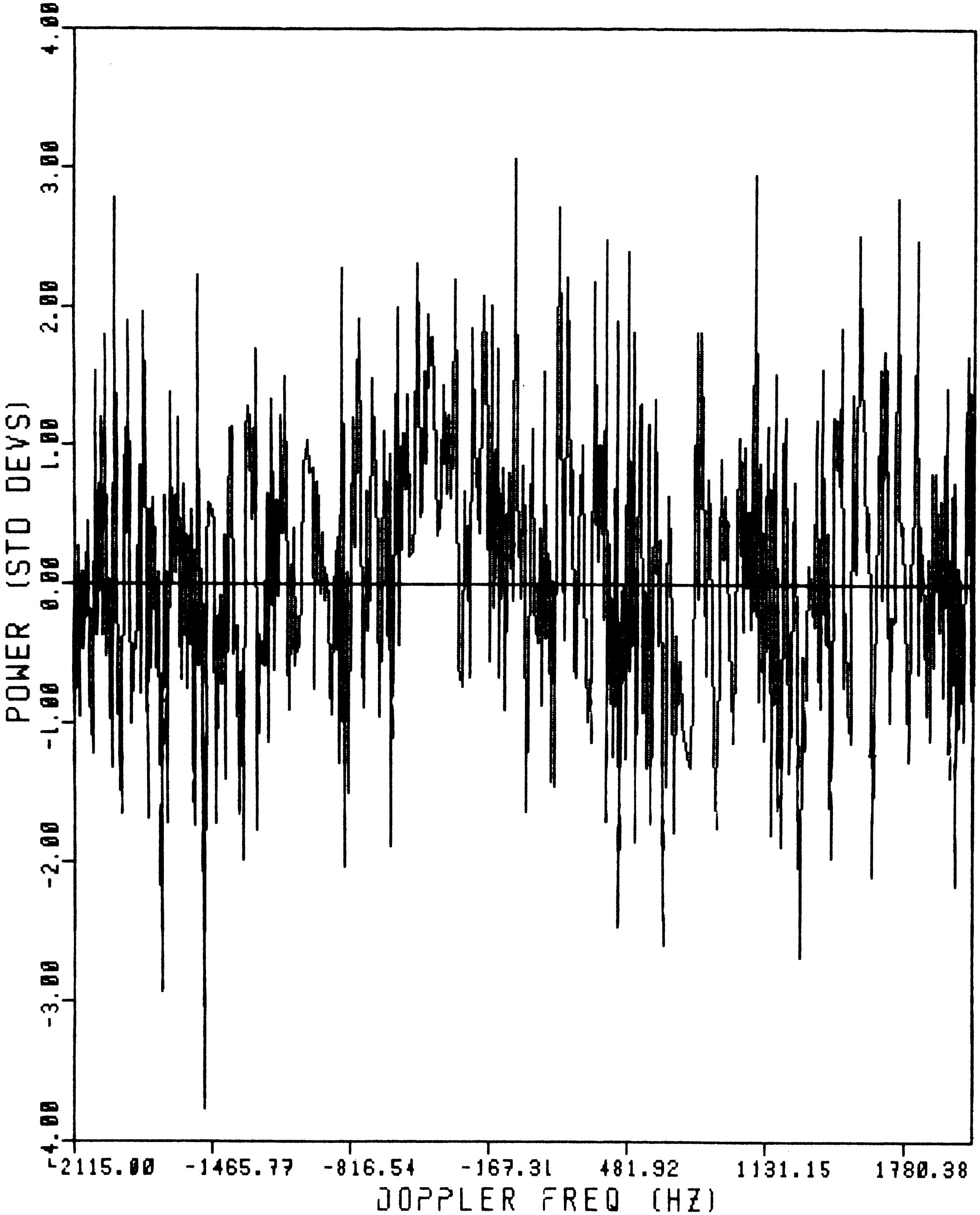

Figure 1. Radar astronomical observations of a minor planet: Echo power spectral density estimates at 433 Doppler frequencies.

$$P[\text{no runs of length } k] \simeq (1-px)/[(k+1-kx)(1-p)x^{n+1}] \quad ,$$

where p is the probability of a 1 at each trial and

$$x = 1 + (1-p)p^k + (k+1)[(1-p)p^k]^2 + \tfrac{1}{2}(k+1)(k+2)[(1-p)p^k]^3 + \cdots$$

is a solution of the equation

$$1 - x + (1-p)p^k x^{k+1} = 0 \quad .$$

It is of interest to examine the accuracy of this approximation.

An exact solution to a more general version of the runs problem, in which the Bernoulli trials are independent but not identically distributed, is obtained in Section 2. The numerics of the problem and the accuracy of Feller's approximation are discussed in Section 3. The poor behavior of the method of Bonferroni inequalities is treated in Section 4.

2. Mathematical formulation and solution. Let $X_1, \cdots, X_n$ be independent and identically distributed with

$$\Pr[X_i = 1] = p, \qquad \Pr[X_i = 0] = 1 - p$$

for $i = 1, 2, \cdots, n$. It is required to determine the probability that there is a run of at least k successive 1's in the sequence $X_1, \cdots, X_n$, where $k \geq 2$. This event occurs if and only if there is some i for which $1 \leq i \leq n-k+1$ and

$$X_i = X_{i+1} = X_{i+2} = \cdots = X_{i+k-1} = 1 \quad .$$

We shall refer to this event as "a run of k" for brevity.

The assumption of identically distributed X_i's, i.e., a common value of p for all trials, is not required for the approach of this section to be applicable. If $\Pr[X_i = 1] = p_i$, where p_i varies with i, the calculation becomes a bit more cumbersome but the method is essentially unchanged. This case will be treated at the end of this section.

There are three parameters: the number n of trials, the length k of the run, and the probability p that any $X_i = 1$. Fix the values of k and p, and define for $m = 1, \cdots, n$

$$r(m) \equiv \Pr[\text{the first run of } k \text{ ends at observation } m] \quad ,$$

$$s(m) \equiv \Pr[\text{there is a run of } k \text{ in the first } m \text{ observations}] \quad .$$

Some initial conditions are easy to establish. There cannot be a run of k in the first m observations when $m < k$, so $r(m) = s(m) = 0$ for $m < k$. For a run of k to occur in the first k observations, all must be 1's, so $r(k) = p^k$. For the first run of k to end at the mth observation where $k + 1 \le m \le 2k$, the first $m - k - 1$ observations can take any values, as there are not enough of them to form a run of k; they must be followed by a zero and then by k successive ones, so $r(m) = (1 - p)p^k$.

Two recursive relations provide $r(m)$ for $m > 2k$ and $s(m)$ for all m. For the first run of k to end at the mth observation where $m > 2k$, there must be no run of k in the first $m - k - 1$ observations, which happens with probability $1 - s(m - k - 1)$, followed by a zero and then by k successive ones, so

$$r(m) = [1 - s(m - k - 1)](1 - p)p^k \quad .$$

When a run of k occurs in the first m observations, there either is or is not a run of k in the first $m - 1$ observations. These situations are disjoint, so the sum of their probabilities is

$$s(m) = s(m - 1) + r(m) \quad .$$

Combining these results gives

$$(1) \qquad r(m) = \begin{cases} 0 & \text{for } m = 1, \cdots, k - 1 \\ p^k & \text{for } m = k \\ p^k(1 - p) & \text{for } m = k + 1, \cdots, 2k \\ [1 - s(m - k - 1)]p^k(1 - p) & \text{for } m = 2k + 1, \cdots, n \end{cases} ,$$

$$(2) \qquad s(m) = \begin{cases} 0 & \text{for } m = 1, \cdots, k-1 \\ s(m-1) + r(m) & \text{for } m = k, \cdots, n \end{cases} \quad .$$

Equations (1) and (2) constitute a recursive solution to the problem. As m increases from k to n, $r(m)$ can be computed for each m from (1) and the known values of $r(j)$ and $s(j)$ with $j = 1, 2, \cdots, m-1$, and $s(m)$ follows easily from (2). The probability of a run of k in a sequence of n Bernoulli trials is given by $s(n)$. It is easy to show that this probability is monotone increasing in n and p, and monotone decreasing in k. Furthermore, for any n and p, Pr[longest run of 1's is exactly k trials] = Pr[a run of k] - Pr[a run of $k+1$] .

Assume now that $X_1, \cdots, X_n$ are independent but not identically distributed, so $\Pr[X_i = 1] = p_i$, where p_i varies with i. In a biological setting, p_i may change with the time of day or day of the week at which a trial occurs. For arbitrary sequences of p_i's, equations (1) and (2) are easily extended to

$$r(m) = \begin{cases} 0 & \text{for } m = 1, \cdots, k-1 \\ \Pi_{i=1}^{k} p_i & \text{for } m = k \\ [1 - p_{m-k}]\Pi_{i=1}^{k} p_{m-k+i} & \text{for } m = k+1, \cdots, 2k \\ [1 - s(m-k-1)][1 - p_{m-k}]\Pi_{i=1}^{k} p_{m-k+i} & \text{for } m = 2k+1, \cdots, n \end{cases} \quad ,$$

$$s(m) = \begin{cases} 0 & \text{for } m = 1, \cdots, k-1 \\ s(m-1) + r(m) & \text{for } m = k, \cdots, n \end{cases} \quad .$$

In this case, the generating function approach of Feller is not applicable, as the regularity it requires is not present.

3. Numerical solution. A FORTRAN program to implement the algorithm of equations (1) and (2) with all calculations done in double precision mode is given below in the appendix. Subroutine PRUN performs the calculation. The main

program provides the necessary input values and formats the output. Table 1 shows the values taken by the probability of a run of k for $n = 50(10)500$, $k = 15(1)30$, when $p = 0.5$.

To test the accuracy of the approximation of Feller given at the end of Section 1, the exact and approximate values of Pr[a run of k] were computed and the percentage of error in the approximation was found for $n = 50(50)500$, $k = 15(1)30$, and $p = 0.2(0.1)0.9$. The approximation appears to behave very well in most of this region. When $p = 0.5$, it is correct to within less than 0.0001% over the entire range of n and k; when $p = 0.6$ and 0.7, it is correct to within 0.0008% and 0.0490%, respectively. For the values $p = 0.8$ and 0.9, the percentage of error in the approximation is shown in Table 2. For large p, the percentage error is greatest when n and k are small, e.g., when $p = 0.9$, $n = 50$, and $k = 15$, the error is -9.35% . For small p, the approximation is not easy to compute with precision, as it is obtained by taking 1 - Pr[no runs of k] . When p is small, k does not have to be very large to make Pr[no runs of k] close enough to 1 to create numerical difficulties even when calculations are made in double precision mode, e.g., when $p = 0.2$, $n = 200$, Pr[a run of 25] $= 0.473 \times 10^{-15}$.

4. A cautionary note on the Bonferroni approach. The Bonferroni inequalities often provide an effective method for obtaining tight upper and lower bounds on a probability that is difficult to compute. The results of applying the Bonferroni approach to the problem considered here indicate that this technique must be used with caution. In particular, the closeness of upper and lower bounds and the convergence of successive bounds cannot be assumed. The details leading to these conclusions constitute the remainder of this section.

Table 1. Probability of a Run of k for p = 0.5, n = 50(10)500, k = 15(1)30

N K:	15	16	17	18
50	0.56452E-03	0.27465E-03	0.13351E-03	0.64850E-04
60	0.71705E-03	0.35093E-03	0.17166E-03	0.83922E-04
70	0.86955E-03	0.42720E-03	0.20980E-03	0.10299E-03
80	0.10220E-02	0.50347E-03	0.24794E-03	0.12207E-03
90	0.11745E-02	0.57973E-03	0.28608E-03	0.14114E-03
100	0.13269E-02	0.65599E-03	0.32422E-03	0.16021E-03
110	0.14793E-02	0.73224E-03	0.36235E-03	0.17928E-03
120	0.16317E-02	0.80849E-03	0.40049E-03	0.19835E-03
130	0.17841E-02	0.88472E-03	0.43862E-03	0.21742E-03
140	0.19364E-02	0.96096E-03	0.47675E-03	0.23649E-03
150	0.20887E-02	0.10372E-02	0.51489E-03	0.25556E-03
160	0.22410E-02	0.11134E-02	0.55301E-03	0.27463E-03
170	0.23933E-02	0.11896E-02	0.59114E-03	0.29370E-03
180	0.25456E-02	0.12658E-02	0.62927E-03	0.31277E-03
190	0.26978E-02	0.13420E-02	0.66739E-03	0.33184E-03
200	0.28500E-02	0.14182E-02	0.70552E-03	0.35090E-03
210	0.30022E-02	0.14944E-02	0.74364E-03	0.36997E-03
220	0.31543E-02	0.15706E-02	0.78176E-03	0.38904E-03
230	0.33064E-02	0.16468E-02	0.81988E-03	0.40810E-03
240	0.34586E-02	0.17230E-02	0.85800E-03	0.42717E-03
250	0.36106E-02	0.17991E-02	0.89611E-03	0.44624E-03
260	0.37627E-02	0.18753E-02	0.93423E-03	0.46530E-03
270	0.39147E-02	0.19515E-02	0.97234E-03	0.48437E-03
280	0.40668E-02	0.20276E-02	0.10105E-02	0.50343E-03
290	0.42187E-02	0.21038E-02	0.10486E-02	0.52250E-03
300	0.43707E-02	0.21799E-02	0.10867E-02	0.54156E-03
310	0.45227E-02	0.22560E-02	0.11248E-02	0.56062E-03
320	0.46746E-02	0.23322E-02	0.11629E-02	0.57969E-03
330	0.48265E-02	0.24083E-02	0.12010E-02	0.59875E-03
340	0.49784E-02	0.24844E-02	0.12391E-02	0.61781E-03
350	0.51302E-02	0.25605E-02	0.12772E-02	0.63688E-03
360	0.52820E-02	0.26366E-02	0.13153E-02	0.65594E-03
370	0.54338E-02	0.27127E-02	0.13534E-02	0.67500E-03
380	0.55856E-02	0.27888E-02	0.13915E-02	0.69406E-03
390	0.57374E-02	0.28649E-02	0.14296E-02	0.71312E-03
400	0.58891E-02	0.29410E-02	0.14677E-02	0.73218E-03
410	0.60408E-02	0.30171E-02	0.15058E-02	0.75124E-03
420	0.61925E-02	0.30931E-02	0.15439E-02	0.77030E-03
430	0.63442E-02	0.31692E-02	0.15820E-02	0.78936E-03
440	0.64958E-02	0.32453E-02	0.16200E-02	0.80842E-03
450	0.66475E-02	0.33213E-02	0.16581E-02	0.82748E-03
460	0.67991E-02	0.33974E-02	0.16962E-02	0.84654E-03
470	0.69506E-02	0.34734E-02	0.17343E-02	0.86559E-03
480	0.71022E-02	0.35494E-02	0.17724E-02	0.88465E-03
490	0.72537E-02	0.36255E-02	0.18105E-02	0.90371E-03
500	0.74052E-02	0.37015E-02	0.18485E-02	0.92277E-03

Table 1 (continued)

N	K: 19	20	21	22
50	0.31471E-04	0.15259E-04	0.73910E-05	0.35763E-05
60	0.41008E-04	0.20027E-04	0.97752E-05	0.47684E-05
70	0.50544E-04	0.24796E-04	0.12159E-04	0.59605E-05
80	0.60081E-04	0.29564E-04	0.14544E-04	0.71526E-05
90	0.69617E-04	0.34332E-04	0.16928E-04	0.83447E-05
100	0.79153E-04	0.39100E-04	0.19312E-04	0.95368E-05
110	0.88690E-04	0.43869E-04	0.21696E-04	0.10729E-04
120	0.98226E-04	0.48637E-04	0.24080E-04	0.11921E-04
130	0.10776E-03	0.53405E-04	0.26464E-04	0.13113E-04
140	0.11730E-03	0.58173E-04	0.28849E-04	0.14305E-04
150	0.12683E-03	0.62941E-04	0.31233E-04	0.15497E-04
160	0.13637E-03	0.67709E-04	0.33617E-04	0.16689E-04
170	0.14590E-03	0.72478E-04	0.36001E-04	0.17881E-04
180	0.15544E-03	0.77246E-04	0.38385E-04	0.19073E-04
190	0.16498E-03	0.82014E-04	0.40769E-04	0.20266E-04
200	0.17451E-03	0.86782E-04	0.43153E-04	0.21458E-04
210	0.18405E-03	0.91550E-04	0.45537E-04	0.22650E-04
220	0.19358E-03	0.96318E-04	0.47921E-04	0.23842E-04
230	0.20312E-03	0.10109E-03	0.50306E-04	0.25034E-04
240	0.21265E-03	0.10585E-03	0.52690E-04	0.26226E-04
250	0.22219E-03	0.11062E-03	0.55074E-04	0.27418E-04
260	0.23172E-03	0.11539E-03	0.57458E-04	0.28610E-04
270	0.24126E-03	0.12016E-03	0.59842E-04	0.29802E-04
280	0.25079E-03	0.12493E-03	0.62226E-04	0.30994E-04
290	0.26032E-03	0.12969E-03	0.64610E-04	0.32186E-04
300	0.26986E-03	0.13446E-03	0.66994E-04	0.33378E-04
310	0.27939E-03	0.13923E-03	0.69378E-04	0.34570E-04
320	0.28893E-03	0.14400E-03	0.71762E-04	0.35762E-04
330	0.29846E-03	0.14876E-03	0.74146E-04	0.36954E-04
340	0.30800E-03	0.15353E-03	0.76530E-04	0.38147E-04
350	0.31753E-03	0.15830E-03	0.78914E-04	0.39339E-04
360	0.32706E-03	0.16307E-03	0.81298E-04	0.40531E-04
370	0.33660E-03	0.16783E-03	0.83682E-04	0.41723E-04
380	0.34613E-03	0.17260E-03	0.86066E-04	0.42915E-04
390	0.35567E-03	0.17737E-03	0.88450E-04	0.44107E-04
400	0.36520E-03	0.18214E-03	0.90834E-04	0.45299E-04
410	0.37473E-03	0.18691E-03	0.93218E-04	0.46491E-04
420	0.38427E-03	0.19167E-03	0.95602E-04	0.47683E-04
430	0.39380E-03	0.19644E-03	0.97986E-04	0.48875E-04
440	0.40333E-03	0.20121E-03	0.10037E-03	0.50067E-04
450	0.41286E-03	0.20598E-03	0.10275E-03	0.51259E-04
460	0.42240E-03	0.21074E-03	0.10514E-03	0.52451E-04
470	0.43193E-03	0.21551E-03	0.10752E-03	0.53643E-04
480	0.44146E-03	0.22028E-03	0.10991E-03	0.54835E-04
490	0.45100E-03	0.22504E-03	0.11229E-03	0.56027E-04
500	0.46053E-03	0.22981E-03	0.11467E-03	0.57219E-04

Table 1 (continued)

N	K: 23	24	25	26
50	0.17285E-05	0.83447E-06	0.40233E-06	0.19372E-06
60	0.23246E-05	0.11325E-05	0.55135E-06	0.26822E-06
70	0.29206E-05	0.14305E-05	0.70036E-06	0.34273E-06
80	0.35167E-05	0.17285E-05	0.84937E-06	0.41723E-06
90	0.41127E-05	0.20266E-05	0.99838E-06	0.49174E-06
100	0.47088E-05	0.23246E-05	0.11474E-05	0.56625E-06
110	0.53048E-05	0.26226E-05	0.12964E-05	0.64075E-06
120	0.59009E-05	0.29206E-05	0.14454E-05	0.71526E-06
130	0.64969E-05	0.32187E-05	0.15944E-05	0.78977E-06
140	0.70930E-05	0.35167E-05	0.17434E-05	0.86427E-06
150	0.76890E-05	0.38147E-05	0.18925E-05	0.93878E-06
160	0.82851E-05	0.41127E-05	0.20415E-05	0.10133E-05
170	0.88811E-05	0.44108E-05	0.21905E-05	0.10878E-05
180	0.94772E-05	0.47088E-05	0.23395E-05	0.11623E-05
190	0.10073E-04	0.50068E-05	0.24885E-05	0.12368E-05
200	0.10669E-04	0.53048E-05	0.26375E-05	0.13113E-05
210	0.11265E-04	0.56029E-05	0.27865E-05	0.13858E-05
220	0.11861E-04	0.59009E-05	0.29355E-05	0.14603E-05
230	0.12457E-04	0.61989E-05	0.30846E-05	0.15348E-05
240	0.13053E-04	0.64969E-05	0.32336E-05	0.16093E-05
250	0.13649E-04	0.67949E-05	0.33826E-05	0.16838E-05
260	0.14246E-04	0.70930E-05	0.35316E-05	0.17583E-05
270	0.14842E-04	0.73910E-05	0.36806E-05	0.18329E-05
280	0.15438E-04	0.76890E-05	0.38296E-05	0.19074E-05
290	0.16034E-04	0.79870E-05	0.39786E-05	0.19819E-05
300	0.16630E-04	0.82851E-05	0.41276E-05	0.20564E-05
310	0.17226E-04	0.85831E-05	0.42766E-05	0.21309E-05
320	0.17822E-04	0.88811E-05	0.44257E-05	0.22054E-05
330	0.18418E-04	0.91791E-05	0.45747E-05	0.22799E-05
340	0.19014E-04	0.94772E-05	0.47237E-05	0.23544E-05
350	0.19610E-04	0.97752E-05	0.48727E-05	0.24289E-05
360	0.20206E-04	0.10073E-04	0.50217E-05	0.25034E-05
370	0.20802E-04	0.10371E-04	0.51707E-05	0.25779E-05
380	0.21398E-04	0.10669E-04	0.53197E-05	0.26524E-05
390	0.21994E-04	0.10967E-04	0.54687E-05	0.27269E-05
400	0.22590E-04	0.11265E-04	0.56178E-05	0.28014E-05
410	0.23186E-04	0.11563E-04	0.57668E-05	0.28759E-05
420	0.23782E-04	0.11861E-04	0.59158E-05	0.29504E-05
430	0.24378E-04	0.12159E-04	0.60648E-05	0.30249E-05
440	0.24974E-04	0.12457E-04	0.62138E-05	0.30995E-05
450	0.25570E-04	0.12755E-04	0.63628E-05	0.31740E-05
460	0.26166E-04	0.13053E-04	0.65118E-05	0.32485E-05
470	0.26762E-04	0.13351E-04	0.66608E-05	0.33230E-05
480	0.27358E-04	0.13649E-04	0.68098E-05	0.33975E-05
490	0.27954E-04	0.13947E-04	0.69589E-05	0.34720E-05
500	0.28550E-04	0.14245E-04	0.71079E-05	0.35465E-05

Table 1 (continued)

N	K: 27	28	29	30
50	0.93133E-07	0.44704E-07	0.21421E-07	0.10245E-07
60	0.13039E-06	0.63330E-07	0.30734E-07	0.14901E-07
70	0.16764E-06	0.81957E-07	0.40047E-07	0.19558E-07
80	0.20489E-06	0.10058E-06	0.49360E-07	0.24215E-07
90	0.24215E-06	0.11921E-06	0.58674E-07	0.28871E-07
100	0.27940E-06	0.13784E-06	0.67987E-07	0.33528E-07
110	0.31665E-06	0.15646E-06	0.77300E-07	0.38184E-07
120	0.35390E-06	0.17509E-06	0.86614E-07	0.42841E-07
130	0.39116E-06	0.19372E-06	0.95927E-07	0.47498E-07
140	0.42841E-06	0.21234E-06	0.10524E-06	0.52154E-07
150	0.46566E-06	0.23097E-06	0.11455E-06	0.56811E-07
160	0.50292E-06	0.24960E-06	0.12387E-06	0.61468E-07
170	0.54017E-06	0.26822E-06	0.13318E-06	0.66124E-07
180	0.57742E-06	0.28685E-06	0.14249E-06	0.70781E-07
190	0.61468E-06	0.30548E-06	0.15181E-06	0.75438E-07
200	0.65193E-06	0.32410E-06	0.16112E-06	0.80094E-07
210	0.68918E-06	0.34273E-06	0.17043E-06	0.84751E-07
220	0.72644E-06	0.36136E-06	0.17975E-06	0.89408E-07
230	0.76369E-06	0.37998E-06	0.18906E-06	0.94064E-07
240	0.80094E-06	0.39861E-06	0.19837E-06	0.98721E-07
250	0.83820E-06	0.41724E-06	0.20769E-06	0.10338E-06
260	0.87545E-06	0.43586E-06	0.21700E-06	0.10803E-06
270	0.91270E-06	0.45449E-06	0.22631E-06	0.11269E-06
280	0.94995E-06	0.47311E-06	0.23563E-06	0.11735E-06
290	0.98721E-06	0.49174E-06	0.24494E-06	0.12200E-06
300	0.10245E-05	0.51037E-06	0.25425E-06	0.12666E-06
310	0.10617E-05	0.52899E-06	0.26357E-06	0.13132E-06
320	0.10990E-05	0.54762E-06	0.27288E-06	0.13597E-06
330	0.11362E-05	0.56625E-06	0.28219E-06	0.14063E-06
340	0.11735E-05	0.58487E-06	0.29151E-06	0.14529E-06
350	0.12107E-05	0.60350E-06	0.30082E-06	0.14994E-06
360	0.12480E-05	0.62213E-06	0.31013E-06	0.15460E-06
370	0.12852E-05	0.64075E-06	0.31945E-06	0.15926E-06
380	0.13225E-05	0.65938E-06	0.32876E-06	0.16391E-06
390	0.13597E-05	0.67801E-06	0.33807E-06	0.16857E-06
400	0.13970E-05	0.69663E-06	0.34739E-06	0.17323E-06
410	0.14342E-05	0.71526E-06	0.35670E-06	0.17788E-06
420	0.14715E-05	0.73389E-06	0.36601E-06	0.18254E-06
430	0.15088E-05	0.75251E-06	0.37533E-06	0.18720E-06
440	0.15460E-05	0.77114E-06	0.38464E-06	0.19185E-06
450	0.15833E-05	0.78977E-06	0.39395E-06	0.19651E-06
460	0.16205E-05	0.80839E-06	0.40327E-06	0.20117E-06
470	0.16578E-05	0.82702E-06	0.41258E-06	0.20582E-06
480	0.16950E-05	0.84565E-06	0.42189E-06	0.21048E-06
490	0.17323E-05	0.86427E-06	0.43121E-06	0.21514E-06
500	0.17695E-05	0.88290E-06	0.44052E-06	0.21979E-06

Table 2. Percentage of Error in Feller Approximation

for $p = 0.8, 0.9$, $n = 50(150)500$, $k = 15(1)30$

p = .8	N	K:	15	16	17	18
	50		-1.2050	-0.8819	-0.6377	-0.4561
	200		-0.6256	-0.5356	-0.4352	-0.3402
	350		-0.2867	-0.2989	-0.2802	-0.2433
	500		-0.1204	-0.1570	-0.1732	-0.1694

	N	K:	19	20	21	22
	50		-0.3238	-0.2290	-0.1616	-0.1140
	200		-0.2583	-0.1919	-0.1402	-0.1010
	350		-0.1997	-0.1573	-0.1201	-0.0895
	500		-0.1517	-0.1274	-0.1021	-0.0789

	N	K:	23	24	25	26
	50		-0.0805	-0.0569	-0.0404	-0.0141
	200		-0.0720	-0.0509	-0.0357	-0.0249
	350		-0.0655	-0.0472	-0.0337	-0.0238
	500		-0.0594	-0.0437	-0.0317	-0.0227

	N	K:	27	28	29	30
	50		-0.0032	-0.0018	-0.0058	-0.0125
	200		-0.0173	-0.0119	-0.0082	-0.0056
	350		-0.0167	-0.0116	-0.0080	-0.0055
	500		-0.0160	-0.0113	-0.0078	-0.0054

p = .9	N	K:	15	16	17	18
	50		-9.3557	-9.0416	-8.5689	-7.9679
	200		-0.3446	-0.5561	-0.8121	-1.0881
	350		-0.0061	-0.0170	-0.0402	-0.0823
	500		-0.0001	-0.0004	-0.0016	-0.0051

	N	K:	19	20	21	22
	50		-7.3073	-6.6362	-5.9851	-5.3719
	200		-1.3554	-1.5882	-1.7686	-1.8879
	350		-0.1485	-0.2394	-0.3505	-0.4724
	500		-0.0134	-0.0299	-0.0582	-0.1006

	N	K:	23	24	25	26
	50		-4.8061	-4.2917	-3.8289	-3.1038
	200		-1.9455	-1.9469	-1.9013	-1.8191
	350		-0.5935	-0.7027	-0.7912	-0.8540
	500		-0.1569	-0.2237	-0.2955	-0.3657

	N	K:	27	28	29	30
	50		-2.5364	-2.1039	-1.7857	-1.5636
	200		-1.7107	-1.5854	-1.4511	-1.3142
	350		-0.8896	-0.8993	-0.8865	-0.8553
	500		-0.4282	-0.4783	-0.5133	-0.5325

For any events $A_1, \cdots, A_N$, the principle of inclusion and exclusion gives

$$\Pr\left[\bigcup_{i=1}^{N} A_i\right] = T_1 - T_2 + T_3 - T_4 \pm \cdots + (-1)^{N-1} T_N \quad , \tag{3}$$

where

$$T_1 = \sum_{i=1}^{N} \Pr[A_i] \ , \qquad T_2 = \sum_{\substack{i,j=1 \\ i<j}}^{N} \Pr[A_i A_j] \ , \qquad T_3 = \sum_{\substack{i,j,\ell=1 \\ i<j<\ell}}^{N} \Pr[A_i A_j A_\ell] \ , \ \cdots \ ,$$

and

$$T_N = \Pr\ [A_1 A_2 \cdots A_N] \quad .$$

The sum of the first t terms on the right-hand side of (3) provides an upper bound to $\Pr[\cup A_i]$ when t is odd and a lower bound when t is even, producing a sequence of Bonferroni inequalities:

$$\begin{aligned} &\Pr[\cup A_i] \le T_1, \qquad \Pr[\cup A_i] \le T_1 - T_2 + T_3, \ \cdots \\ &T_1 - T_2 \le \Pr[\cup A_i], \qquad T_1 - T_2 + T_3 - T_4 \le \Pr[\cup A_i], \ \cdots \quad . \end{aligned} \tag{4}$$

In many cases, these bounds increase in sharpness as the number of terms included from the right-hand side increases, and this property is often assumed to hold. It will now be shown that this property fails to hold in the situation of this paper.

Let n, k, and p be fixed. Define the events

$$A \equiv \{\text{there is a run of } k\} \quad ,$$

and

$$A_i \equiv \{\text{there is a run of } k \text{ starting at } i\} \qquad \text{for } i = 1, 2, \cdots, n-k+1 \quad ,$$

so $A = \cup_{i=1}^{n-k+1} A_i$. The terms T_1, T_2, T_3, and T_4 of (3) will now be considered, where N denotes $n-k+1$. Since $\Pr[A_i] = p^k$ for each $i = 1, \cdots, N$, it is immediate that

$$T_1 = \sum_{i=1}^{N} \Pr[A_i] = Np^k \quad . \tag{5}$$

Let $\Sigma_{[d]}$ denote summation over the set of index pairs $\{(i,j): 1 \le i < j \le N,\ j-i=d\}$. If $P[d] = p^{k+\min(d,k)}$ denotes the common value of $\Pr[A_iA_j]$ for all pairs (i,j) in this set and $\#[d]$ denotes the number of such pairs, then

$$\sum_{[d]} \Pr[A_iA_j] = \#[d]P[d] = \begin{cases} (N-d)p^{k+d} & \text{for } d = 1, 2, \cdots, k-1 \ , \\ (N-d)p^{2k} & \text{for } d = k, \cdots, n-k \ . \end{cases}$$

Summing these terms over d gives

$$\begin{aligned} T_2 &= \sum_{d=1}^{k-1} \sum_{[d]} \Pr[A_iA_j] + \sum_{d=k}^{n-k} \sum_{[d]} \Pr[A_iA_j] \\ &= \sum_{d=1}^{k-1} (N-d)p^{k+d} + \sum_{d=k}^{n-k} (N-d)p^{2k} = Np^k[p+p^2+p^3+\cdots+p^{k-1}] \\ &\quad - p^k[p+2p^2+3p^3+\cdots+(k-1)p^{k-1}] + p^{2k}[(N-k)+(N-k-1)+\cdots+1] \\ &= Np^kG(0) - p^kG(1) + \tfrac{1}{2}(N-k)(N-k+1)p^{2k} \end{aligned} \tag{6}$$

where

$$G(m) = \sum_{j=1}^{k-1} j^m p^j = \begin{cases} (p-p^k)/(1-p) & \text{for } m=0 \ , \\ [p-kp^k+(k-1)p^{k+1}]/(1-p)^2 & \text{for } m=1 \ , \\ [p+p^2-k^2p^k+(2k^2-2k-1)p^{k+1}-(k-1)^2p^{k+2}]/(1-p)^3 & \text{for } m=2 \ . \end{cases} \tag{7}$$

For calculating T_3, let $\Sigma_{[d,e]}$ denote summation over the set of index triples $\{(i,j,\ell): 1 \le i < j < \ell \le N,\ j-i=d,\ \ell-j=e\}$. The number of triples in this set is easily seen to be

$$\#[d,e] = N-d-e \quad ,$$

and for all of these triples,

$$\Pr[A_i A_j A_\ell] \equiv P[d,e] = p^{k+\min(d,k)+\min(e,k)} \quad .$$

Summing $\Sigma_{[d,e]}\Pr[A_i A_j A_\ell]$ over all pairs $\{[d,e]: 1 \le d,e \le N-2;\ d+e \le N-1\}$ gives

$$
\begin{aligned}
T_3 &= \sum_{[d,e]} \sum\nolimits_{[d,e]} \Pr[A_i A_j A_\ell] \\
&= \sum_{[d,e]} \#[d,e] P[d,e] \qquad (8) \\
&= \sum_{[d,e]} (N-d-e) p^{k+\min(d,k)+\min(e,k)} \quad .
\end{aligned}
$$

To evaluate this sum, group the pairs $[d,e]$ by the exponent of p in $P[d,e]$ as in Table 3, where

$$
\begin{aligned}
[d,k\!\uparrow] &\equiv \{[d,k],\ [d,k+1],\ \cdots,\ [d,N-d-1]\} \quad , \\
[k\!\uparrow,e] &\equiv \{[k,e],\ [k+1,e],\ \cdots,\ [N-e-1,e]\} \quad , \\
[k\!\uparrow,k\!\uparrow] &\equiv [k,k\!\uparrow] \cup [k+1,k\!\uparrow] \cup [k+2,k\!\uparrow] \cup \cdots \cup [N-k-1,k\!\uparrow] \quad .
\end{aligned}
$$

It is straightforward to show that

$$
\begin{aligned}
\#[d,k\!\uparrow] &= \sum_{e=k}^{N-d-1} \#[d,e] = \tfrac{1}{2}(N-k-d)(N-k-d+1) \quad , \\
\#[k\!\uparrow,e] &= \sum_{d=k}^{N-e-1} \#[d,e] = \tfrac{1}{2}(N-k-e)(N-k-e+1) \quad , \\
\#[k\!\uparrow,k\!\uparrow] &= \sum_{d=k}^{N-k-1} \#[d,k\!\uparrow] = \tfrac{1}{2} \sum_{i=1}^{N-2k} i(i+1) \\
&= (N-2k)(N-2k+1)(N-2k+2)/6 \quad .
\end{aligned}
$$

Combining these results with Table 3 and equations (8) and (7) gives

Table 3. Exponents of p in P[d,e] and Corresponding [d,e] Patterns

Exponent of p	Corresponding [d,e] Patterns
k+2	[1,1]
k+3	[1,2] [2,1]
k+4	[1,3] [2,2] [3,1]
k+5	[1,4] [2,3] [3,2] [4,1]
⋮	⋮
k+j	$\underbrace{[1,j-1]\ [2,j-2] \cdots [j-1,1]}_{j-1 \text{ terms}}$
⋮	⋮
2k	[1,k-1] [2,k-2] ··· [k-1,1]
2k+1	$[1,k\uparrow]\ [k\uparrow,1]\ [2,k-1]\ [3,k-2] \cdots [k-1,2]$
2k+2	$[2,k\uparrow]\ [k\uparrow,2]\ [3,k-1]\ [4,k-2] \cdots [k-1,3]$
⋮	⋮
2k+j	$[j,k\uparrow]\ [k\uparrow,j]\ \underbrace{[j+1,k-1]\ [j+2,k-2] \cdots [k-1,j+1]}_{k-j-1 \text{ terms}}$
⋮	⋮
3k-1	$[k-1,k\uparrow]\ [k\uparrow,k-1]$
3k	$[k\uparrow,k\uparrow]$

$$T_3 = \sum_{j=2}^{k} (j-1)(N-j)p^{k+j} + \sum_{j=1}^{k-1} [(N-k-j)(N-k-j+1) + (k-j-1)(N-k-j)]p^{2k+j}$$

(9) $$+ (N-2k)(N-2k+1)(N-2k+2)p^{3k}/6$$

$$= p^{k+1}[(N-1)G(1) - G(2)] + p^{2k}[N(N-k)G(0) - (3N-2k)G(1) + 2G(2)]$$

$$+ (N-2k)(N-2k+1)(N-2k+2)p^{3k}/6 \quad .$$

Finally, using notation analogous to that just developed for T_3,

$$T_4 = \sum_{[d,e,f]} \#[d,e,f]P[d,e,f]$$

(10)

$$= \sum_{[d,e,f]} (N-d-e-f)p^{k+\min(d,k)+\min(e,k)+\min(f,k)} \quad .$$

Although T_4 could be computed by defining patterns $[d,e,k\uparrow]$, etc., constructing a table like Table 3, and substituting into the formula just given, the present problem does not require that this cumbersome calculation be fully treated. Instead, it suffices to obtain a lower bound L_4 for T_4 by summing the terms of equation (10) over all $[d,e,f]$ with $d+e+f \equiv j \leq k+1$. For any such $[d,e,f]$, the exponent of p in (10) is $k+j$, since $d,e,f < k$, and $N-d-e-f = N-j$, so

$$T_4 \geq L_4 = \sum_{\substack{[d,e,f] \\ j\equiv d+e+f\leq k+1}} (N-d-e-f)p^{k+\min(d,k)+\min(e,k)+\min(f,k)}$$

$$= \sum_{j=3}^{k+1} \mathcal{T}(j) \times (N-j)p^{k+j} \quad ,$$

where $\mathcal{T}(j)$ denotes the number of triples $[d,e,f]$ of positive integers with $d+e+f = j$. This is just the number of distinct ordered triples of non-negative integers whose three elements sum to $j - 3$, or equivalently, the number of distinct arrangements of $j - 3$ 0's and two 1's, so $\mathcal{J}(j) = \frac{1}{2}(j-1)(j-2)$, and

$$(11) \qquad L_4 = \sum_{j=3}^{k+1} \tfrac{1}{2}(j-1)(j-2)(N-j)p^{k+j}$$

$$= \tfrac{1}{2}p^{k+2}[(N-2)G(1) + (N-3)G(2) - G(3)] \quad .$$

Intuitively, this analysis has examined the first segment of a table analogous to Table 3, with exponent of p ranging from $k+3$ to $2k+1$. For exponent $k+j$, the number of corresponding $[d,e,f]$ patterns is $\mathcal{J}(j)$. These patterns could be listed as in Table 3, e.g., for $j=6$, $\mathcal{J}(j)=10$, and the list consists of three permutations of $[1,1,4]$, six permutations of $[1,2,3]$, and $[2,2,2]$. The number of distinct sets of four event indices i, j, ℓ, h with $j-i=d$, $\ell-j=e$, $h-\ell=f$ is $\#[d,e,f] = N-j$, and p^{k+j} is $\Pr[A_iA_jA_\ell A_h]$ for any of these sets.

The first three Bonferroni bounds of (4) are obtained easily from equations (5), (6), (9), and (11). Tables 4 and 5 compare the first upper Bonferroni bound T_1, the first lower Bonferroni bound $T_1 - T_2$, the second upper Bonferroni bound $T_1 - T_2 + T_3$, and the upper limit $T_1 - T_2 + T_3 - L_4$ for the second lower Bonferroni bound to some exact values of $\Pr[\cup A_i]$. The ratios of these Bonferroni bounds to the corresponding exact probabilities appear in Table 5.

The first upper Bonferroni bound is larger than the exact probability by a factor ranging from 1.97 to 4.53, and the second upper Bonferroni bound is larger than the exact probability by a factor ranging from 1.92 to 39. Neither of these bounds is close to the exact probability for any combination of n, k, and p, and the second upper bound is often greater than the first. The first lower Bonferroni bound is never more than .055 of the exact probability, and

Table 4. First Three Bonferroni Bounds and an Upper Limit for the Fourth Bonferroni Bound to P[run of length k] for p = 0.5(0.1)0.7, k = 15(5)30, n = 100(100)500

N	K	P	T1-T2+T3-L4	T1-T2	PR(K RUN)	T1	T1-T2+T3
100	15	0.5	0.1284E-03	0.5878E-04	0.1327E-02	0.2625E-02	0.2565E-02
200	15	0.5	0.1455E-03	0.4765E-04	0.2850E-02	0.5676E-02	0.5628E-02
300	15	0.5	0.1720E-03	0.2721E-04	0.4371E-02	0.8728E-02	0.8700E-02
400	15	0.5	0.2079E-03	-0.2548E-05	0.5889E-02	0.1178E-01	0.1178E-01
500	15	0.5	0.2532E-03	-0.4162E-04	0.7405E-02	0.1483E-01	0.1487E-01
100	20	0.5	0.3822E-05	0.1906E-05	0.3910E-04	0.7725E-04	0.7534E-04
200	20	0.5	0.3842E-05	0.1896E-05	0.8678E-04	0.1726E-03	0.1707E-03
300	20	0.5	0.3871E-05	0.1877E-05	0.1345E-03	0.2680E-03	0.2661E-03
400	20	0.5	0.3910E-05	0.1848E-05	0.1821E-03	0.3634E-03	0.3615E-03
500	20	0.5	0.3957E-05	0.1811E-05	0.2298E-03	0.4587E-03	0.4569E-03
100	25	0.5	0.1192E-06	0.5960E-07	0.1147E-05	0.2265E-05	0.2205E-05
200	25	0.5	0.1192E-06	0.5959E-07	0.2638E-05	0.5245E-05	0.5186E-05
300	25	0.5	0.1193E-06	0.5957E-07	0.4128E-05	0.8225E-05	0.8166E-05
400	25	0.5	0.1193E-06	0.5955E-07	0.5618E-05	0.1121E-04	0.1115E-04
500	25	0.5	0.1194E-06	0.5951E-07	0.7108E-05	0.1419E-04	0.1413E-04
100	30	0.5	0.3725E-08	0.1863E-08	0.3353E-07	0.6612E-07	0.6426E-07
200	30	0.5	0.3725E-08	0.1863E-08	0.8009E-07	0.1593E-06	0.1574E-06
300	30	0.5	0.3725E-08	0.1863E-08	0.1267E-06	0.2524E-06	0.2505E-06
400	30	0.5	0.3725E-08	0.1862E-08	0.1732E-06	0.3455E-06	0.3437E-06
500	30	0.5	0.3725E-08	0.1862E-08	0.2198E-06	0.4387E-06	0.4368E-06
100	15	0.6	-0.5444E-01	-0.1898E-01	0.1637E-01	0.4044E-01	0.6818E-01
200	15	0.6	-0.1228E+00	-0.4512E-01	0.3475E-01	0.8745E-01	0.1556E+00
300	15	0.6	-0.1865E+00	-0.7347E-01	0.5278E-01	0.1345E+00	0.2476E+00
400	15	0.6	-0.2456E+00	-0.1040E+00	0.7048E-01	0.1815E+00	0.3443E+00
500	15	0.6	-0.2999E+00	-0.1368E+00	0.8785E-01	0.2285E+00	0.4458E+00
100	20	0.6	-0.4140E-02	-0.1346E-02	0.1206E-02	0.2962E-02	0.4913E-02
200	20	0.6	-0.1002E-01	-0.3189E-02	0.2666E-02	0.6618E-02	0.1134E-01
300	20	0.6	-0.1588E-01	-0.5045E-02	0.4124E-02	0.1027E-01	0.1779E-01
400	20	0.6	-0.2171E-01	-0.6914E-02	0.5580E-02	0.1393E-01	0.2427E-01
500	20	0.6	-0.2751E-01	-0.8797E-02	0.7034E-02	0.1759E-01	0.3078E-01
100	25	0.6	-0.3003E-03	-0.9738E-04	0.8813E-04	0.2161E-03	0.3568E-03
200	25	0.6	-0.7619E-03	-0.2396E-03	0.2018E-03	0.5004E-03	0.8545E-03
300	25	0.6	-0.1223E-02	-0.3819E-03	0.3155E-03	0.7847E-03	0.1352E-02
400	25	0.6	-0.1685E-02	-0.5243E-03	0.4292E-03	0.1069E-02	0.1850E-02
500	25	0.6	-0.2146E-02	-0.6668E-03	0.5429E-03	0.1353E-02	0.2349E-02
100	30	0.6	-0.2157E-04	-0.7019E-05	0.6411E-05	0.1570E-04	0.2581E-04
200	30	0.6	-0.5749E-04	-0.1807E-04	0.1525E-04	0.3780E-04	0.6450E-04
300	30	0.6	-0.9341E-04	-0.2913E-04	0.2410E-04	0.5991E-04	0.1032E-03
400	30	0.6	-0.1293E-03	-0.4018E-04	0.3294E-04	0.8202E-04	0.1419E-03
500	30	0.6	-0.1652E-03	-0.5124E-04	0.4178E-04	0.1041E-03	0.1806E-03
100	15	0.7	-0.2475E+01	-0.5599E+00	0.1205E+00	0.4083E+00	0.1715E+01
200	15	0.7	-0.4955E+01	-0.1459E+01	0.2398E+00	0.8830E+00	0.4667E+01
300	15	0.7	-0.6448E+01	-0.2584E+01	0.3428E+00	0.1358E+01	0.8606E+01
400	15	0.7	-0.6847E+01	-0.3934E+01	0.4319E+00	0.1833E+01	0.1364E+02
500	15	0.7	-0.6044E+01	-0.5510E+01	0.5089E+00	0.2307E+01	0.1987E+02
100	20	0.7	-0.4580E+00	-0.8105E-01	0.1984E-01	0.6463E-01	0.2464E+00
200	20	0.7	-0.1092E+01	-0.1943E+00	0.4313E-01	0.1444E+00	0.5989E+00
300	20	0.7	-0.1701E+01	-0.3140E+00	0.6588E-01	0.2242E+00	0.9754E+00
400	20	0.7	-0.2286E+01	-0.4400E+00	0.8808E-01	0.3040E+00	0.1376E+01
500	20	0.7	-0.2846E+01	-0.5724E+00	0.1098E+00	0.3838E+00	0.1803E+01
100	25	0.7	-0.7377E-01	-0.1257E-01	0.3149E-02	0.1019E-01	0.3814E-01
200	25	0.7	-0.1872E+00	-0.3063E-01	0.7156E-02	0.2360E-01	0.9389E-01
300	25	0.7	-0.3000E+00	-0.4886E-01	0.1115E-01	0.3701E-01	0.1503E+00
400	25	0.7	-0.4122E+00	-0.6728E-01	0.1512E-01	0.5042E-01	0.2074E+00
500	25	0.7	-0.5236E+00	-0.8588E-01	0.1908E-01	0.6384E-01	0.2651E+00
100	30	0.7	-0.1151E-01	-0.1959E-02	0.4958E-03	0.1600E-02	0.5937E-02
200	30	0.7	-0.3082E-01	-0.4969E-02	0.1172E-02	0.3854E-02	0.1522E-01
300	30	0.7	-0.5011E-01	-0.7983E-02	0.1847E-02	0.6108E-02	0.2452E-01
400	30	0.7	-0.6938E-01	-0.1100E-01	0.2522E-02	0.8362E-02	0.3384E-01
500	30	0.7	-0.8863E-01	-0.1403E-01	0.3196E-02	0.1062E-01	0.4318E-01

Table 5. Ratios of First Three Bonferroni Bounds and an Upper Limit for the Fourth to P[run of length k] for p = 0.5(0.1)0.7, k = 15(5)30, n = 100(100)500

N	K	P	LOWER2/PR	LOWER1/PR	UPPER1/PR	UPPER2/PR
100	15	0.5	0.09676	0.04430	1.97790	1.93341
200	15	0.5	0.05107	0.01672	1.99169	1.97476
300	15	0.5	0.03936	0.00623	1.99694	1.99050
400	15	0.5	0.03531	-0.00043	2.00027	2.00050
500	15	0.5	0.03420	-0.00562	2.00285	2.00830
100	20	0.5	0.09775	0.04874	1.97563	1.92689
200	20	0.5	0.04427	0.02184	1.98908	1.96723
300	20	0.5	0.02879	0.01396	1.99302	1.97906
400	20	0.5	0.02147	0.01015	1.99493	1.98477
500	20	0.5	0.01722	0.00788	1.99606	1.98817
100	25	0.5	0.10390	0.05195	1.97403	1.92208
200	25	0.5	0.04521	0.02259	1.98870	1.96611
300	25	0.5	0.02889	0.01443	1.99278	1.97835
400	25	0.5	0.02124	0.01060	1.99470	1.98410
500	25	0.5	0.01679	0.00837	1.99581	1.98744
100	30	0.5	0.11111	0.05555	1.97222	1.91667
200	30	0.5	0.04651	0.02325	1.98837	1.96512
300	30	0.5	0.02941	0.01470	1.99265	1.97794
400	30	0.5	0.02150	0.01075	1.99462	1.98387
500	30	0.5	0.01695	0.00847	1.99576	1.98729
100	15	0.6	-3.32652	-1.15988	2.47087	4.16626
200	15	0.6	-3.53374	-1.29865	2.51704	4.47886
300	15	0.6	-3.53423	-1.39201	2.54773	4.69169
400	15	0.6	-3.48492	-1.47606	2.57505	4.88550
500	15	0.6	-3.41391	-1.55728	2.60115	5.07482
100	20	0.6	-3.43277	-1.11595	2.45538	4.07296
200	20	0.6	-3.75972	-1.19595	2.48206	4.25286
300	20	0.6	-3.85078	-1.22320	2.49115	4.31417
400	20	0.6	-3.89096	-1.23909	2.49644	4.34993
500	20	0.6	-3.91186	-1.25068	2.50030	4.37603
100	25	0.6	-3.40773	-1.10497	2.45166	4.04866
200	25	0.6	-3.77479	-1.18715	2.47906	4.23356
300	25	0.6	-3.87696	-1.21041	2.48681	4.28589
400	25	0.6	-3.92474	-1.22158	2.49053	4.31100
500	25	0.6	-3.95231	-1.22824	2.49275	4.32600
100	30	0.6	-3.36410	-1.09484	2.44828	4.02588
200	30	0.6	-3.76879	-1.18482	2.47827	4.22835
300	30	0.6	-3.87643	-1.20879	2.48626	4.28227
400	30	0.6	-3.92626	-1.21990	2.48997	4.30728
500	30	0.6	-3.95498	-1.22633	2.49211	4.32174
100	15	0.7	-20.53670	-4.64552	3.38752	14.22799
200	15	0.7	-20.66898	-6.08652	3.68314	19.46402
300	15	0.7	-18.80995	-7.53759	3.96073	25.10262
400	15	0.7	-15.85280	-9.10878	4.24294	31.57835
500	15	0.7	-11.87648	-10.82620	4.53374	39.05046
100	20	0.7	-23.08786	-4.08610	3.25841	12.42060
200	20	0.7	-25.30812	-4.50529	3.34835	13.88432
300	20	0.7	-25.82460	-4.76603	3.40358	14.80658
400	20	0.7	-25.95738	-4.99524	3.45152	15.62769
500	20	0.7	-25.93369	-5.21485	3.49688	16.42390
100	25	0.7	-23.42509	-3.99059	3.23636	12.11096
200	25	0.7	-26.16542	-4.27971	3.29840	13.12125
300	25	0.7	-26.91805	-4.38371	3.32069	13.48507
400	25	0.7	-27.25882	-4.44947	3.33477	13.71551
500	25	0.7	-27.44576	-4.50105	3.34578	13.89659
100	30	0.7	-23.21953	-3.95059	3.22755	11.97460
200	30	0.7	-26.30577	-4.24087	3.28977	12.99026
300	30	0.7	-27.13075	-4.32261	3.30729	13.27623
400	30	0.7	-27.51154	-4.36338	3.31603	13.41889
500	30	0.7	-27.72966	-4.38915	3.32155	13.50909

the upper limit for the second lower Bonferroni bound is never more than .111 of the exact probability. Both lower Bonferroni bounds are often negative, and the second lower bound is less than the first in many cases.

The bad behavior of the Bonferroni bounds is due to the highly dependent nature of the events A_i . Tight bounds result when some T_t is very small. The quantities $T_1, T_2, T_3, \cdots$ consist of $n, \binom{n}{2}, \binom{n}{3}, \cdots$ terms, each of which is a probability. Therefore T_2 cannot be small relative to T_1 unless all of its terms are small relative to terms in T_1 . This does not happen, since

$$P[A_i A_{i+1}] = p\, P[A_i] \qquad \text{for } i = 1, \cdots, N-1$$

and

$$P[A_i A_{i+2}] = p^2 P[A_i] \qquad \text{for } i = 1, \cdots, N-2 \quad ,$$

and so on. Summing these terms shows that T_2 is of the same order of magnitude as T_1 . The same line of reasoning applies to T_3 and higher-order quantities.

References

Feller, W. (1968). An Introduction to Probability Theory and Its Applications, Vol. I, 3rd edition, John Wiley and Sons, Inc., New York.

Gibbons, J. D. (1971). Nonparametric Statistical Inference, McGraw-Hill Book Company, New York.

Mood, A. M. (1940). The distribution theory of runs. Ann. Math. Statist. 11, 367-392.

Mosteller, F. (1941). Note on an application of runs to quality control charts. Ann. Math. Statist. 12, 228-232.

Ostro, S. J., Campbell, D. B., and Shapiro, I. I. (1980). Private communication.

Shaughnessy, P. W. (1981). Multiple runs distributions: recurrences and critical values. JASA 76, 732-736.

Appendix

```
C     THIS PROGRAM CALCULATES PROBABILITY OF A RUN OF K (OR MORE)
C       SUCCESSES IN A SEQUENCE OF N IID BERNOULLI TRIALS
      DOUBLE PRECISION R(500), S(500), P
      DIMENSION PROB(46,16), N(46)
      COMMON R,S,P,K
  110 FORMAT('1PROBABILITY OF RUN OF LENGTH K WHEN P(SUCCESS) =',F5.1/
     C    1X,55('-')//4X,'N   K:',5X,'15',12X,'16',12X,'17',12X,'18'/
     D    1X,64('-')/(1X,I4,3X,4E14.5))
  120 FORMAT('1PROBABILITY OF RUN OF LENGTH K WHEN P(SUCCESS) =',F5.1/
     C    1X,55('-')//4X,'N   K:',5X,'19',12X,'20',12X,'21',12X,'22'/
     D    1X,64('-')/(1X,I4,3X,4E14.5))
  130 FORMAT('1PROBABILITY OF RUN OF LENGTH K WHEN P(SUCCESS) =',F5.1/
     C    1X,55('-')//4X,'N   K:',5X,'23',12X,'24',12X,'25',12X,'26'/
     D    1X,64('-')/(1X,I4,3X,4E14.5))
  140 FORMAT('1PROBABILITY OF RUN OF LENGTH K WHEN P(SUCCESS) =',F5.1/
     C    1X,55('-')//4X,'N   K:',5X,'27',12X,'28',12X,'29',12X,'30'/
     D    1X,64('-')/(1X,I4,3X,4E14.5))
  160 FORMAT(5E16.8)
      DO 200 I=1,46
         N(I) = 10*I + 40
  200 CONTINUE
      DO 500 L=5,7
         P = .1*FLOAT(L)
         DO 400 K=15,30
            CALL PRUN
            KM14 = K - 14
            DO 300 I=1,46
               PROB(I,KM14) = S(10*I + 40)
  300       CONTINUE
  400    CONTINUE
         WRITE(6,110) P,(N(I),(PROB(I,K),K=1,4),I=1,46)
         WRITE(6,120) P,(N(I),(PROB(I,K),K=5,8),I=1,46)
         WRITE(6,130) P,(N(I),(PROB(I,K),K=9,12),I=1,46)
         WRITE(6,140) P,(N(I),(PROB(I,K),K=13,16),I=1,46)
         WRITE(9,160) ((PROB(I,K),I=6,46,10),K=1,16,5)
  500 CONTINUE
      STOP
      END
C
      SUBROUTINE PRUN
      DOUBLE PRECISION R(500), S(500), P, PK, PK1P, D
      COMMON R,S,P,K
      PK   = P**K
      PK1P = PK*(1.0D0 - P)
      R(K) = PK
      S(K) = PK
      KP1  = K + 1
      KK   = 2*K
      DO 200 M=KP1,KK
         R(M) = PK1P
         D    = FLOAT(M-K)
         S(M) = PK + D*PK1P
  200 CONTINUE
      KKP1 = KK + 1
      DO 300 M=KKP1,500
         MK1  = (M - K) - 1
         R(M) = (1.0D0 - S(MK1))*PK1P
         S(M) = S(M-1) + R(M)
  300 CONTINUE
      RETURN
      END
```

Appendix (continued)

```
C     THIS PROGRAM COMPUTES THE FIRST THREE BONFERRONI BOUNDS AND AN
C       UPPER LIMIT FOR THE FOURTH BONFERRONI BOUND TO PR(RUN OF
C       LENGTH K), AND RATIOS OF THESE BOUNDS TO PR(RUN OF LENGTH K),
C       FOR P=0.5(0.1)0.7, K=15(5)30, N=100(100)500
      DOUBLE PRECISION P, Q, PK, XK, XLN, XBN, G0,G1,G2, T1,T2,T3, XL4
      DIMENSION PROB(60), IRATIO(60,2), RATIO(60,5)
  100 FORMAT(5E16.8)
  110 FORMAT('1   N     K    P',2X,'T1-T2+T3-L4',4X,'T1-T2',5X,
     C     'PR(K RUN)',6X,'T1',7X,'T1-T2+T3'/1X,75('-'))
  120 FORMAT(1X,2I4,F5.1,5E12.4)
  130 FORMAT('1   N     K    P     LOWER2/PR  LOWER1/PR  UPPER1/PR   ',
     C     'UPPER2/PR'/1X,59('-')/(1X,2I4,F5.1,4F11.5))
      READ(9,100) PROB
      KT = 1
      WRITE(6,110)
      DO 600 L=5,7
        P = 0.1*FLOAT(L)
        Q = 1.0D0 - P
        DO 500 K=15,30,5
          PK = P**K
          XK = K
          KP1 = K+1
          DO 400 N=100,500,100
C           XLN = LITTLE N;  XBN = NBIG = BIG N = (LITTLE N) - K + 1
            XLN = N
            NBIG = N + 1 - K
            XBN = NBIG
            G0 = (P - PK)/Q
            G1 = (P + (XK-1.D0)*PK*P - XK*PK)/Q**2
            G2 = (P + P*P + (2.D0*XK**2 - (2.D0*XK + 1.D0))*PK*P -
     C             (XK**2*PK + (XK-1.D0)**2*PK*P**2))/Q**3
            T1 = XBN*PK
            T2 = T1*G0 + 0.5D0*(XBN-XK)*(XBN+1.D0-XK)*PK**2 - PK*G1
            T3 = ((XBN-1.D0)*G1 - G2)*PK*P + (XBN*(XBN-XK)*G0
     C             - (3.D0*XBN-2.D0*XK)*G1 + 2.D0*G2)*PK**2
     D             + (XBN-2.D0*XK)*(XBN+1.D0-2.D0*XK)
     E             *(XBN+2.D0-2.D0*XK)*PK**3/6.D0
            XL4 = 0.D0
            DO 300 J=3,KP1
              XL4 = XL4 + P**J*DBLE(FLOAT((J-1)*(J-2)*(NBIG-J)/2))
  300       CONTINUE
            XL4 = XL4*PK
            B1 = T1
            B2 = T1 - T2
            B3 = T1 - T2 + T3
            B4 = T1 - T2 + T3 - XL4
            WRITE(6,120) N,K,P,B4,B2,PROB(KT),B1,B3
            IRATIO(KT,1) = N
            IRATIO(KT,2) = K
            RATIO(KT,1)  = P
            RATIO(KT,2)  = B4/PROB(KT)
            RATIO(KT,3)  = B2/PROB(KT)
            RATIO(KT,4)  = B1/PROB(KT)
            RATIO(KT,5)  = B3/PROB(KT)
            KT = KT + 1
  400     CONTINUE
  500   CONTINUE
  600 CONTINUE
      WRITE(6,130) ((IRATIO(KT,M),M=1,2),(RATIO(KT,M),M=1,5),KT=1,60)
      STOP
      END
```